

From Microsoft Agent to Operator AI

An Evolution of Agentic
Architecture from COM to
LangChain

Legacy MS Agent vs Operator Workflow

- Legacy (Microsoft Agent):
 - - spchapi.dll (SAPI based TTS)
 - - COM/OLE Dispatch
 - - ActiveX controls (e.g., Merlin, Genie)
 - - Local .ACS scripts and animations
- Modern (Operator / LangChain):
 - - TTS APIs (OpenAI, ElevenLabs, Azure)

Speech Layer: From spchapi to Whisper

- Microsoft SAPI stack:
 - - spchapi.dll handles TTS and SR interfaces
 - - Requires x86 ActiveX environment
- Modern Layer:
 - - OpenAI TTS (server-side, streaming)
 - - Whisper for ASR (Automatic Speech Recognition)
 - - Connected via LangChain/Agent

System Skeleton: Modules and Kernels

- - systemd boot and unit orchestration
- - Operator kernel loads browser headlessly
- - FastAPI + LangChain combo acts as orchestrator
- - Modular design mirrors init layers from Debian

Interactive UX + Spectrograms

